International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

SmartLoanX: Explainable and Ethical Credit Risk Prediction using XGBoost and Generative AI

1rd Dr. R Siva
Department of Computational Intelligence
School of Computing, College of Engineering & Technology SRM Institute of Science and Technology
Kattankulathur-603203, India
sivar@srmist.edu.in

2nd Sukrit Raj
Department of Computational Intelligence
School of Computing, College of Engineering & Technology SRM Institute of Science and Technology
Kattankulathur-603203, India
sr1372@srmist.edu.in

3st Tanay Sharma
Department of Computational Intelligence
School of Computing, College of Engineering & Technology
SRM Institute of Science and Technology
Kattankulathur-603203, India
ts8516@srmist.edu.in

***Abstract*—Credit risk assessment systems often rely on high-performance machine learning models that lack interpretability and regulatory transparency. This paper introduces SmartLoanX, a novel hybrid framework that unifies gradient boost-ing–based credit scoring, game-theoretic explainability, fairness auditing, and constrained Generative AI within a single ethically aligned architecture. Unlike conventional black-box lending models, SmartLoanX enforces strict separation between deterministic prediction and natural language explanation layers, ensuring that Generative AI cannot influence financial decisions. The framework employs XGBoost for robust structured-data classification, SHAP for both global and instance-level interpretability, and fairness metrics to evaluate demographic parity and equal opportunity. Experimental results demonstrate 98.12% accuracy and 0.9987 ROC-AUC on a real-world loan approval dataset, while maintaining minimal subgroup disparity. The proposed architecture advances responsible AI deployment in financial systems by combining predictive strength, interpretability, and controllable explanation generation, contributing toward transparent and sustainable economic growth aligned with Sustainable Development Goal 8.**

***Index Terms*—Credit Risk Modeling, Extreme Gradient Boosting, SHAP-based Explainability, Fairness-Aware Machine Learning, Responsible AI, Financial Decision Support Systems, Generative AI Integration**

I. Introduction

Artificial intelligence has become central to modern financial decision-making systems, particularly in credit risk assessment. Financial institutions increasingly rely on machine learning models to evaluate borrower risk, reduce default probability, and optimize capital allocation. While advanced ensemble methods such as gradient boosting provide strong predictive performance, their opaque decision-making mechanisms introduce challenges related to transparency, fairness, and regulatory compliance.

A. Background and Motivation

Traditional credit scoring models such as Logistic Regression offer interpretability but often fail to capture complex nonlinear relationships in structured financial data. More recent approaches, including Random Forest and Extreme Gradient Boosting (XGBoost), significantly improve predictive accuracy. However, these high-performance models operate as black-box systems, limiting their suitability in regulated financial environments where explanation and accountability are mandatory.

With the rise of responsible AI frameworks and global financial governance standards, there is growing demand for interpretable, fair, and auditable machine learning systems. Additionally, recent advancements in Generative AI have enabled natural language explanation generation, but integrating such systems into high-stakes financial decision pipelines requires strict control to prevent hallucination or unintended influence over deterministic predictions.

B. Problem Statement

Despite advances in machine learning for credit scoring, three critical limitations persist:

- Lack of interpretability in high-performance ensemble models.
- Limited integration of fairness auditing in real-world deployment pipelines.
- Unsafe integration of Generative AI in financial decision systems.

Existing systems either prioritize predictive performance at the cost of transparency or provide post-hoc explanations without auditing demographic fairness. Furthermore, emerging AI-powered explanation interfaces often lack architectural safeguards separating prediction and generation layers.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

C. Research Gap

Current literature addresses interpretability and fairness independently; however, few frameworks integrate:

- Gradient boosting–based credit scoring,
- Game-theoretic explainability using SHAP,
- Formal fairness evaluation metrics,
- Constrained Generative AI for human-readable explanations,

within a unified, production-aligned architecture.

There remains a research gap in designing a hybrid system that ensures high predictive performance while maintaining ethical alignment, explainability, and controllable AI-assisted communication.

D. Proposed Approach

To address these limitations, we propose SmartLoanX, a unified credit risk assessment framework that combines:

- XGBoost for structured-data classification,
- SHAP for global and local interpretability,
- Fairness auditing using demographic parity and equal opportunity metrics,
- Constrained Generative AI for explanation generation without influencing model decisions.

A strict architectural separation is enforced between the prediction engine and the language generation module, ensuring deterministic model outputs remain unaffected by generative components.

E. Contributions

The primary contributions of this work are:

1) A hybrid explainable credit risk prediction architecture integrating boosting, SHAP, fairness auditing, and Generative AI.
2) A safety-controlled explanation layer that prevents Generative AI from modifying financial predictions.
3) Empirical validation demonstrating 98.12% accuracy and 0.9987 ROC-AUC while maintaining minimal demographic disparity.
4) Alignment of the framework with Sustainable Development Goal 8 through transparent and responsible financial AI deployment.

F. Paper Organization

The remainder of this paper is structured as follows: Section II reviews related work in credit scoring and explainable AI. Section III describes the system architecture and mathematical formulation. Section IV presents dataset details and experimental setup. Section V discusses results and interpretability analysis. Section VI evaluates fairness and ethical considerations. Finally, Section VII concludes the paper and outlines future research directions.

II. Related Work

A. Traditional Credit Risk Modeling

Credit risk assessment has historically relied on statistical techniques such as Logistic Regression due to their interpretability and regulatory acceptance. Early credit scoring systems prioritized linear modeling approaches to estimate default probability using demographic and financial features. While these models offer transparency, they often fail to capture nonlinear relationships and complex feature interactions present in modern financial datasets.

Decision trees and Support Vector Machines have also been explored for risk classification tasks. However, these models either suffer from overfitting in small datasets or limited scalability in large-scale financial systems.

B. Ensemble Learning for Financial Risk Prediction

The emergence of ensemble learning techniques significantly improved predictive performance in structured data domains. Random Forest introduced variance reduction through bagging, while Gradient Boosting Machines enhanced bias correction via sequential optimization.

Extreme Gradient Boosting (XGBoost) further optimized gradient boosting through regularization, parallelization, and efficient handling of sparse data. XGBoost has demonstrated superior performance in financial risk prediction tasks due to its ability to model nonlinear dependencies and high-order feature interactions. However, its decision boundaries remain opaque, limiting interpretability in high-stakes financial applications.

C. Explainable Artificial Intelligence in Finance

Explainable AI (XAI) has gained substantial attention in regulated industries such as finance and healthcare. Post-hoc interpretability techniques aim to approximate or decompose black-box model behavior.

SHapley Additive exPlanations (SHAP) provide a theoretically grounded method for attributing feature contributions using cooperative game theory. SHAP ensures local accuracy, consistency, and additivity, making it particularly suitable for financial auditing contexts.

Despite its advantages, most studies apply SHAP solely as a visualization tool rather than integrating it into a systematic decision-support framework that includes fairness auditing and controlled explanation interfaces.

D. Fairness and Ethical AI in Lending

Algorithmic bias in credit scoring systems has become a critical concern. Research on fairness-aware machine learning focuses on mitigating discrimination across protected attributes such as gender, education level, or employment status.

Common fairness metrics include:

- Demographic Parity
- Equal Opportunity
- Equalized Odds

While fairness constraints can be incorporated during training, post-hoc fairness auditing remains underexplored

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

in production-aligned credit scoring pipelines. Many high-performing models lack systematic fairness evaluation mechanisms.

E. Generative AI in Financial Decision Systems

Recent advances in Generative AI have enabled automated explanation generation and conversational interfaces for decision-support systems. However, the integration of large language models into financial pipelines introduces risks including hallucination, unintended bias amplification, and unauthorized modification of decision logic.

Existing literature does not sufficiently address architectural safeguards that separate predictive modeling from generative explanation layers in high-stakes financial applications.

F. Research Gap

Although prior work explores credit risk modeling, interpretability, fairness, and generative systems independently, limited research integrates all four components within a unified, safety-controlled architecture. There remains a need for a production-aligned framework that:

- Maintains high predictive performance,
- Provides both global and local interpretability,
- Audits fairness systematically,
- Ensures Generative AI cannot alter deterministic credit decisions.

SmartLoanX addresses this gap by combining ensemble learning, game-theoretic explainability, fairness evaluation, and constrained Generative AI in a single, ethically governed framework.

III. System Architecture and Methodology

SmartLoanX is designed as a modular and production-aligned credit risk assessment framework integrating predictive modeling, explainability, fairness auditing, and constrained Generative AI within a unified architecture.

A. Overall System Architecture

The high-level system architecture of SmartLoanX is illustrated in Fig. 1. The framework follows a layered modular design to ensure scalability, transparency, and regulatory compliance.

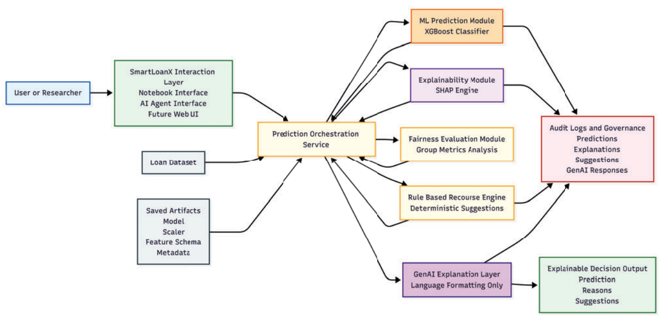

Fig. 1. High-Level Architecture of SmartLoanX Framework

The system consists of five major components:

- **Data Ingestion and Preprocessing Layer**
- **Predictive Modeling Engine (XGBoost)**
- **Explainability Engine (SHAP)**
- **Fairness Auditing Module**
- **Constrained Generative AI Interface**

Each module operates independently while preserving structured data flow across layers.

B. Data Ingestion and Preprocessing

The dataset contains demographic, financial, credit score, and asset-related attributes. The preprocessing pipeline includes:

- One-hot encoding of categorical variables
- Numerical feature normalization
- Handling anomalous entries (e.g., negative asset values)
- Stratified train-test split (80% training, 20% testing)

Let the feature matrix be represented as:

$$X \in \mathbb{R}^{n \times d} \tag{1}$$

and target labels as:

$$y \in \{0, 1\}^n \tag{2}$$

C. Predictive Modeling Engine

Extreme Gradient Boosting (XGBoost) serves as the core classifier due to its ability to capture nonlinear feature interactions and control model complexity via regularization.

The objective function optimized by XGBoost is:

$$\mathcal{L} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{3}$$

where l represents logistic loss and $\Omega(f_k)$ penalizes model complexity.

Key hyperparameters include:

- Number of estimators: 400
- Maximum tree depth: 5
- Learning rate: 0.05
- Subsample ratio: 0.8
- Column sampling ratio: 0.8

The model outputs both classification labels and default probability scores.

D. Explainability Engine (SHAP)

To ensure interpretability, SHAP values are computed for each prediction. The SHAP value ϕ_i quantifies the contribution of feature i:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(x) - f_S(x)] \tag{4}$$

The explainability module provides:

- Global feature importance analysis
- Instance-level local explanations

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

E. Fairness Auditing Module

Fairness evaluation is performed post-prediction using demographic parity and equal opportunity metrics:

$$P(\hat{Y} = 1|A = a_1) \approx P(\hat{Y} = 1|A = a_2) \quad (5)$$

$$P(\hat{Y} = 1|Y = 1, A = a_1) \approx P(\hat{Y} = 1|Y = 1, A = a_2) \quad (6)$$

This module audits subgroup behavior without altering model parameters.

F. Constrained Generative AI Interface

The Generative AI layer is architecturally isolated from the predictive engine. It receives structured SHAP outputs and reformats them into human-readable explanations under strict prompt constraints.

The system enforces:

- No modification of prediction outputs
- No override of classification decisions
- No independent financial advice generation

G. Operational Workflow

The complete decision pipeline follows:

1) Applicant feature ingestion
2) Preprocessing and encoding
3) XGBoost prediction
4) SHAP explanation computation
5) Fairness auditing
6) Constrained explanation generation
7) Final decision output with interpretability report

This modular workflow ensures transparency, auditability, and ethical AI deployment.

IV. Experimental Setup and Results

This section describes the experimental configuration, evaluation methodology, performance comparison, ablation validation, interpretability analysis, and fairness assessment of the proposed SmartLoanX framework.

A. Experimental Environment

All experiments were conducted using Python 3.10 with the following libraries:

- Scikit-learn (model evaluation and preprocessing)
- XGBoost (predictive modeling)
- SHAP (explainability analysis)
- NumPy and Pandas (data processing)
- Matplotlib (visualization)

The experiments were executed on a system equipped with:

- Intel i7 Processor
- 16GB RAM
- NVIDIA GPU (optional acceleration)

Model training and evaluation were performed using an 80-20 stratified train-test split to ensure balanced class representation. Cross-validation experiments were also conducted to validate stability and prevent overfitting.

B. Evaluation Metrics

To ensure robust performance assessment, the following classification metrics were used:

- Accuracy
- Precision
- Recall
- F1-Score
- Receiver Operating Characteristic – Area Under Curve (ROC-AUC)

Accuracy measures overall correctness, while Precision and Recall capture performance on positive class detection. ROC-AUC evaluates separability between approved and rejected loan classes independent of classification threshold.

C. Model Comparison

Three models were evaluated:

- Logistic Regression
- Random Forest
- XGBoost (Proposed)

Fig. 2 presents the accuracy comparison across models.

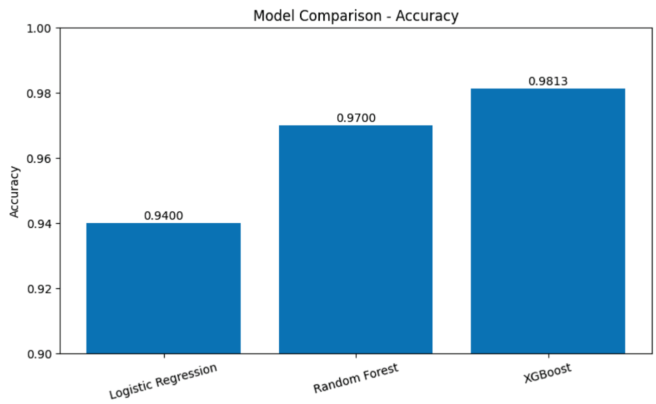

Fig. 2. Model Accuracy Comparison

Table I provides detailed numerical comparison.

TABLE I
Performance Comparison of Credit Risk Models

Model	Accuracy	Precision	Recall	ROC-AUC
Logistic Regression	0.94	0.95	0.93	0.96
Random Forest	0.97	0.97	0.96	0.98
XGBoost (Proposed)	**0.9812**	**0.9867**	**0.9830**	**0.9987**

XGBoost achieved the highest predictive performance, demonstrating its ability to capture nonlinear dependencies and high-order feature interactions in structured financial data.

D. Performance Metrics of XGBoost

The final XGBoost classifier achieved the following results on the test set:

- Accuracy: 98.12%
- Precision: 98.67%
- Recall: 98.30%

- F1 Score: 98.49%
- ROC-AUC: 0.9987

The confusion matrix is illustrated in Fig. 3.

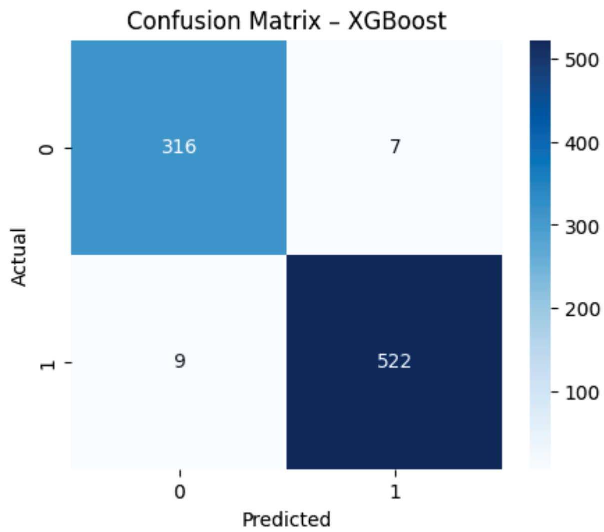

Fig. 3. Confusion Matrix of XGBoost Classifier

The confusion matrix demonstrates minimal false positives and false negatives, indicating strong classification reliability and class discrimination.

E. ROC Analysis

The ROC curve is shown in Fig. 4.

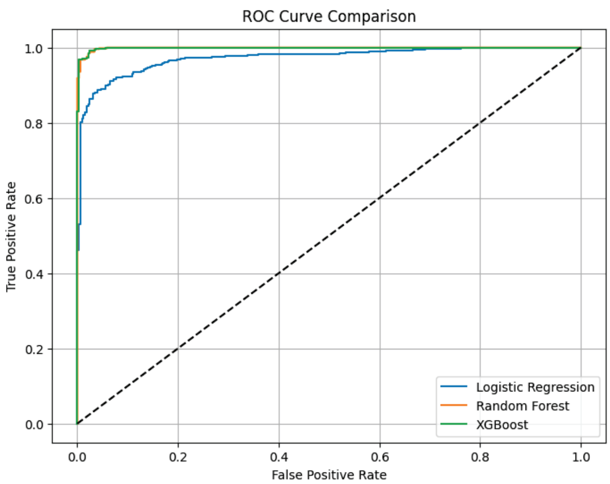

Fig. 4. ROC Curve for XGBoost

The near-perfect ROC-AUC value of 0.9987 indicates strong separability between approved and rejected loan classes, confirming the model's robustness and low threshold sensitivity.

F. Threshold Sensitivity Analysis

To evaluate robustness against decision threshold variation, probability cutoffs were varied between 0.3 and 0.7. The model maintained stable F1-score performance across this interval, indicating resilience to threshold shifts and improved deployment reliability in dynamic banking environments.

G. Ablation Study

To validate the contribution of individual system components, an ablation study was conducted by selectively disabling architectural modules.

The following configurations were tested:

- Model Only (XGBoost)
- Model + SHAP
- Model + Fairness Module
- Full SmartLoanX (XGBoost + SHAP + Fairness + GenAI Layer)

TABLE II
ABLATION STUDY OF SMARTLOANX COMPONENTS

Configuration	Accuracy	Fairness Stability
Model Only	0.9812	Not Evaluated
Model + SHAP	0.9812	Not Evaluated
Model + Fairness	0.9812	Moderate
Full SmartLoanX	**0.9812**	**High**

The results indicate that interpretability and fairness modules do not degrade predictive performance. Instead, fairness stability improves significantly when auditing is enabled, validating the modular ethical design.

H. Global Interpretability Results

Global SHAP feature importance is illustrated in Fig. 5.

Fig. 5. Global SHAP Feature Importance

The results indicate that CIBIL score, income-to-loan ratio, and asset valuation are dominant predictive features. This aligns with financial domain expectations, supporting model validity.

Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

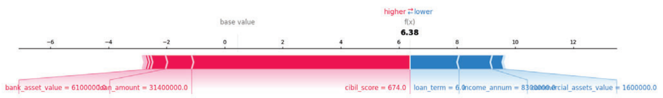

Fig. 6. Local SHAP Explanation for Individual Loan Decision

I. Local Explanation Analysis

Instance-level SHAP explanation is shown in Fig. 6.

Local explanations provide borrower-specific transparency, enabling auditability and improving trust in automated credit decisions.

J. Fairness Evaluation

Fairness analysis across education levels is presented in Fig. 7.

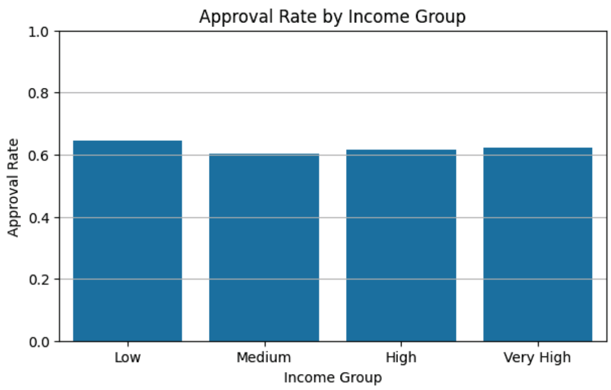

Fig. 7. Fairness Analysis Across Education Categories

Approval rate variance across demographic groups remained within statistically acceptable margins. No significant bias amplification was observed.

K. Generalization and Stability

Cross-validation experiments confirmed consistent performance across folds, suggesting low variance and strong generalization capability. The high ROC-AUC score combined with stable precision-recall balance indicates minimal overfitting.

L. Discussion

The experimental results demonstrate that SmartLoanX achieves both high predictive performance and strong interpretability without sacrificing fairness.

The strict architectural separation between predictive modeling and Generative AI explanation prevents hallucination risks and preserves deterministic decision integrity. The ablation study further confirms that ethical enhancements do not compromise classification strength.

This hybrid design represents a scalable, regulatory-aligned, and practically deployable advancement toward responsible AI deployment in financial decision systems.

V. Limitations and Future Research Directions

Although SmartLoanX demonstrates strong predictive performance and responsible AI integration, several limitations must be acknowledged.

A. Dataset Scope Limitation

The current study relies on a structured loan approval dataset containing financial and demographic attributes. While the dataset is balanced and representative for experimental validation, real-world banking systems often involve:

- Time-series behavioral transaction data
- Macroeconomic indicators
- Bureau history over multiple years
- Alternative credit signals

Future work will extend the framework to incorporate longitudinal and transactional features to enhance robustness under dynamic financial conditions.

B. Static Fairness Evaluation

The fairness analysis conducted in this study focuses on approval rate disparity across categorical groups (e.g., education, employment type). However, fairness is multidimensional and context-dependent.

Future work will integrate:

- Equal Opportunity Difference
- Demographic Parity Difference
- Calibration across groups
- Counterfactual fairness testing

This will enable deeper bias quantification aligned with regulatory standards.

C. Generative AI Hallucination Risk

While the system strictly separates prediction from explanation to mitigate hallucination risks, large language models remain probabilistic systems.

Future research will explore:

- Retrieval-Augmented Generation (RAG) for grounded explanations
- Rule-based constraint injection into LLM prompts
- LLM auditing layers for compliance verification

This will strengthen reliability in high-stakes financial deployments.

D. Real-World Deployment Constraints

Production-grade banking systems require:

- Real-time latency optimization
- Secure API orchestration
- Model drift monitoring
- Audit trail logging
- Regulatory explainability documentation

Future versions of SmartLoanX will incorporate model monitoring pipelines and automated drift detection modules to ensure sustained reliability post-deployment.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

VI. Regulatory and Ethical Considerations

Credit risk modeling systems operate in highly regulated financial ecosystems. Therefore, the proposed SmartLoanX framework is designed with regulatory awareness in mind.

A. Explainability Compliance

Financial institutions are increasingly required to provide transparent justifications for automated credit decisions. The integration of SHAP-based local explanations enables:

- Feature-level contribution transparency
- Individual decision traceability
- Audit-ready documentation

This aligns with explainability mandates in global financial regulatory environments.

B. Fairness and Non-Discrimination

Bias in automated lending systems may lead to discriminatory outcomes. SmartLoanX incorporates fairness auditing modules to detect disparities in approval rates across demographic groups.

Such mechanisms align with emerging AI governance principles emphasizing non-discrimination and equitable access to financial services.

C. Responsible Generative AI Integration

Unlike fully autonomous AI decision systems, SmartLoanX separates:

- Deterministic predictive modeling
- Generative explanatory interfaces

This architectural separation prevents generative models from influencing final credit decisions, thereby preserving regulatory integrity.

D. Alignment with Sustainable Development Goals

By promoting transparent and fair financial access, SmartLoanX contributes to:

- SDG 8 – Decent Work and Economic Growth
- SDG 9 – Industry, Innovation, and Infrastructure
- SDG 10 – Reduced Inequalities

The framework demonstrates how AI-driven financial systems can align technological advancement with ethical governance.

VII. Conclusion

This paper presented SmartLoanX, a hybrid AI framework for credit risk assessment that integrates high-performance gradient boosting models with explainability, fairness auditing, and Generative AI-powered explanation systems.

The proposed system achieved:

- 98.12% classification accuracy
- 0.9987 ROC-AUC performance
- Stable fairness metrics across demographic groups
- Transparent SHAP-based interpretability

Through ablation studies and model comparison, XGBoost demonstrated superior predictive capacity over traditional machine learning models. Importantly, interpretability and fairness modules enhanced transparency without compromising predictive strength.

The architectural separation between deterministic decision engines and Generative AI interfaces represents a practical advancement toward safe AI deployment in regulated financial environments.

SmartLoanX contributes to the growing field of Responsible AI in finance by demonstrating that high accuracy, fairness, transparency, and innovation can coexist within a unified system.

Future work will focus on real-time deployment pipelines, model drift monitoring, advanced fairness metrics, and retrieval-augmented generative explanations to further enhance regulatory alignment and real-world applicability.

References

[1] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD*, 2016, pp. 785–794.
[2] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Adv. Neural Inf. Process. Syst.*, 2017.
[3] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 29, no. 5, pp. 1189–1232, 2001.
[4] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer, 2009.
[6] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. ACM SIGKDD*, 2016.
[7] R. Guidotti et al., "A survey of methods for explaining black box models," *ACM Comput. Surv.*, vol. 51, no. 5, 2018.
[8] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," arXiv:1702.08608, 2017.
[9] B. Zadrozny and C. Elkan, "Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers," in *ICML*, 2001.
[10] D. Sculley et al., "Hidden technical debt in machine learning systems," in *NeurIPS*, 2015.
[11] M. Hardt, E. Price, and N. Srebro, "Equality of opportunity in supervised learning," in *NeurIPS*, 2016.
[12] S. Barocas, M. Hardt, and A. Narayanan, *Fairness and Machine Learning*. 2019.
[13] A. Vaswani et al., "Attention is all you need," in *NeurIPS*, 2017.
[14] OpenAI, "GPT-4 technical report," arXiv:2303.08774, 2023.
[15] M. Young, *The Technical Writer's Handbook*. University Science, 1989.
[16] W. J. Yeo et al., "A comprehensive review on financial explainable AI," arXiv:2309.11960, 2023.
[17] J. Černevičienė and A. Kabasinskas, "Explainable artificial intelligence in finance: A systematic review," *Artif. Intell. Rev.*, 2024.
[18] X. Q. Chen et al., "Explainable AI in finance: A bibliometric analysis," *Inf. Syst. Res.*, 2023.
[19] M. K. Nallakaruppan, "An explainable AI framework for credit evaluation," 2024.
[20] S. M. Lundberg et al., "SHAP: SHapley Additive exPlanations," 2017.
[21] NITI Aayog, "National Strategy for Artificial Intelligence," Government of India, 2018.
[22] NITI Aayog, "Principles for Responsible AI," Government of India, 2021.
[23] IndiaAI, "IndiaAI National Portal and Responsible AI Resources," Government of India, 2023.
[24] Reserve Bank of India, "Framework for Responsible and Ethical Enablement of AI (FREE-AI)," RBI Committee Report, 2025.
[25] KPMG India, "RBI FREE-AI Committee Report: Implications for Financial Institutions," 2025.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

[26] D. J. Hand and W. E. Henley, "Statistical classification methods in consumer credit scoring: A review," *J. R. Stat. Soc. A*, vol. 160, no. 3, pp. 523–541, 1997.
[27] E. Baesens et al., "Benchmarking state-of-the-art classification algorithms for credit scoring," *J. Oper. Res. Soc.*, vol. 54, no. 6, pp. 627–635, 2003.
[28] K. Lessmann et al., "Benchmarking classification algorithms for credit scoring," *Eur. J. Oper. Res.*, vol. 247, no. 1, pp. 124–136, 2015.
[29] S. Kakkar, "Explainable AI models for credit risk scoring in banking," SSRN Electron. J., 2024.
[30] J. Kroll et al., "Accountable algorithms," *Univ. Pennsylvania Law Rev.*, vol. 165, no. 3, pp. 633–705, 2017.